New condition for invariance of ellipsoidal sets for discrete-time saturated systems

Rachid Riah, Mirko Fiacchini

New condition for invariance of ellipsoidal sets for discrete-time saturated systems

Rachid Riah[1] and Mirko Fiacchini[1]

Abstract— **In this paper, we consider the problem of characterizing the invariant and contractive ellipsoids for discrete-time saturated systems, as an estimate of the domain of attraction. The asymptotic stability of the controlled system is ensured by the contractive nature of the invariant set. Sufficient conditions for the existence of a quadratic set-induced Lyapunov function are established through the solution of a bilinear matrix inequalities (BMI) problem. Some computational considerations are analyzed to overcome the problem of complexity. An illustrative example is given to demonstrate the effectiveness of the proposed results, developed in this paper.**

Kaywords: Invariance, Contractivity, Lyapunov function, Convex set, Saturated system.

I. INTRODUCTION

In the last decade, the stability of linear systems subject to actuator saturation attracted a big interest. There are many researchers working on this topic currently, see [1], [4], [8], [12], [13], [16], [19]. The maximal output admissible sets have important applications in the analysis and design of closed-loop systems with state and control constraints [10]. Moreover, input saturation is a nonlinearity encountered in many industrial application because, in general, the actuators can not provide signals that exceed their capacity [18]. The linear systems may not be globally asymptotically stabilized by linear feedback, when subject to input saturations [4], [11], [18], [20]. Lyapunov theory permits to analyse the stability and convergence of such systems. One possible computationally suitable approach is based on the search of Lyapunov functions induced by invariant and contractive sets. See [5] for more details about invariance. Therefore, one important application of invariant and contractive sets is to analyse the stability and convergence properties of such systems. The work presented in this paper deals with the discrete-time systems with input saturations.

In this paper, we consider the problem of characterizing and computing contractive sets for discrete-time saturated systems. This sets provide estimations of the domain of attraction and determine set-induced Lyapunov function, see [1], [2], [8]. As a popular candidate invariant set, the ellipsoids have been widely used as an estimation of the domain of attraction for such systems, see [1], [8]. Sufficient conditions are given for an ellipsoid to be invariant or contractive. They allow to characterize the level set of a local quadratic Lyapunov function, see [1], [8], [15]. The work in this paper is the extension of the results in [9], specifying the convex set as an ellipsoid.

[1] GIPSA-lab, Grenoble Campus, 11 rue des Mathématiques, BP 46, 38402 Saint Martin d'Hères Cedex, France. {rachid.riah, mirko.fiacchini}@gipsa-lab.grenoble-inp.fr

As the results are an invariant contractive ellipsoid and a quadratic Lyapunov function, this permits to compare them with similar approaches, see [1], to demonstrate the effectiveness of the given results. The illustrative example shows that the presented sufficient conditions are less conservative than those found in the literature.

The paper is organized as follows. In Section II, the problem statement is introduced. Some preliminary results are given in Section III. The main results are stated in the Section IV, where the sufficient conditions for the contractivity of a given ellipsoid are presented. The computational considerations are analyzed in the Section V. The comparison with other methods and a discussion on the obtained results are also given in the Section V. Section VI finishes the paper with some conclusions and future works.

Notation: Given $\mathbb{R}$, define $\mathbb{R}_+$ as the non negative real numbers. Given $n \in \mathbb{N}$, define $\mathbb{N}_n = \{x \in \mathbb{N} : 1 \leq x \leq n\}$. Given $A \in \mathbb{R}^{n\times m}$, A_i with $i \in \mathbb{N}_n$ denotes its i-th row, $A_{(j)}$ with $j \subseteq \mathbb{N}_m$ its j-th column. A C-set is a convex, compact set $\Omega \subseteq \mathbb{R}^n$ with $0 \in \text{int}(\Omega)$. Given the C-set Ω and $\alpha \geq 0$, define the set $\alpha\Omega = \{\alpha x \in \mathbb{R}^n : x \in \Omega\}$, the interior of Ω is $\text{int}(\Omega)$ and its boundary is $\partial\Omega$. Given $J \subseteq \mathbb{N}_n$, define $\bar{J}$ as the complement of J in $\mathbb{N}_n$. Given $u \in \mathbb{R}^n$, define $\|u\|_2 = (u^T u)^{1/2}$ as the euclidean norm of u.

II. PROBLEM STATEMENT

Consider the saturated discrete-time linear system given by

$$x^+ = f(x) = Ax + B\varphi(Kx), \tag{1}$$

where $x \in \mathbb{R}^n$ is the current state, $x^+ \in \mathbb{R}^n$ is the successor and the saturated feedback control is given by $u = \varphi(Kx) \in \mathbb{R}^m$. Function $\varphi : \mathbb{R}^m \to \mathbb{R}^m$ denotes the saturation function which is defined by $\varphi_i(y) = \text{sgn}(y_i)\min\{|y_i|, 1\}$ for every $i \in \mathbb{N}_m$.

The main aim of this paper is to establish an approach to compute the contractively invariant ellipsoidal set as an estimation of the domain of attraction for this class of systems. Let now introduce some useful tools to deal with convex closed sets.

Definition 1: Given a set $\Omega \subseteq \mathbb{R}^n$, the support function of Ω evaluated at $\eta \in \mathbb{R}^n$ is $\phi_\Omega(\eta) = \sup_{x\in\Omega} \eta^T x$.

Geometrically, the support function of Ω at η is the signed "distance" of the point of the closure of Ω farthest from the origin, along the direction η. See [17] for some properties of support functions.

Using the support function defined above, the set-inclusion conditions can be given in terms of linear inequalities, as recalled here, see [17], for instance.

Property 1: Given two closed, convex sets $\Omega \subseteq \mathbb{R}^n$ and $\Gamma \subseteq \mathbb{R}^n$, then $x \in \Omega$ if and only if $\eta^T x \le \phi_\Omega(\eta)$ for all $\eta \in \mathbb{R}^n$, and $\Gamma \subseteq \Omega$ if and only if $\phi_\Gamma(\eta) \le \phi_\Omega(\eta)$, for all $\eta \in \mathbb{R}^n$.

Two important concepts that will be used in this paper are the invariance and λ-contractivity of sets. Below, we give some definitions.

Definition 2 ([6]): A C-set $\Omega \subseteq \mathbb{R}^n$ is an invariant set for the system $x^+ = f(x)$ if $f(x) \in \Omega$, for all $x \in \Omega$.

Every trajectory starting in an invariant Ω remains in it.

Definition 3 ([6]): A C-set $\Omega \subseteq \mathbb{R}^n$ is a λ-contractive set for the system $x^+ = f(x)$ if $f(x) \in \lambda\Omega$, for all $x \in \Omega$, with $\lambda \in [0, 1]$.

Since λ-contractivity implies invariance, when in the following we will guarantee λ-contractivity, we will implicitly ensure also invariance. The property of λ-contractivity of a compact convex set can be used to induce a local Lyapunov function. We are interested here to give conditions on compact convex sets $\Omega \subseteq \mathbb{R}^n$, with $0 \in \text{int}(\Omega)$, whose satisfaction ensures that every set $\alpha\Omega$, with $\alpha \in [0, 1]$, is λ-contractive. This would imply that there exists a local Lyapunov function defined on Ω, whose level sets are $\alpha\Omega$ with $\alpha \in [0, 1]$.

Given a C-set Ω , the Minkowski or gauge function of Ω at x is defined as

$$\Psi_\Omega(x) = \min_{\alpha \ge 0}\{\alpha \in \mathbb{R} : \, x \in \alpha\Omega\}.$$

Intuitively, the value of $\Psi_\Omega(x)$ is how much the set Ω should be scaled for x to be on its boundary, that is such that $x \in \partial(\Psi_\Omega(x)\,\Omega)$. Then $x \in \partial\Omega(x)$ where we define

$$\Omega(x) = \Psi_\Omega(x)\Omega. \tag{2}$$

The set $\Omega(x)$ is useful to determine the condition for the sets $\alpha\Omega$ to be λ-contractive for the saturated system (1), for all $\alpha \in [0, 1]$. Such a condition is given by a (possibly uncountable) set of nonconvex constraints, as stated below.

Proposition 1: Given the system (1), the C-set Ω is such that $\alpha\Omega$ is λ-contractive for every $\alpha \in [0, 1]$ if and only if

$$\eta^T f(x) \le \lambda \phi_{\Omega(x)}(\eta), \tag{3}$$

for all $x \in \Omega$ and every $\eta \in \mathbb{R}^n$.

Proof: Sets $\alpha\Omega$ are λ-contractive for every $\alpha \in [0, 1]$ if and only if $x^+ \in \lambda\Omega(x)$, for all $x \in \Omega$. This is equivalent, by Property 1, to (3) for every $x \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^n$. ■

The λ-contractivity of sets $\alpha\Omega$ is equivalent to the decreasing of $\Psi_\Omega(x)$, then it implies local asymptotic stability of the system (1).

III. PRELIMINARY RESULTS

To present the main result of the paper, the following results about how to find the maximal contractively invariant sets for discrete-time saturated systems given in [9], are introduced in this section, where necessary and sufficient conditions for invariance of convex sets are established.

The objective of this work is to render these conditions computationally tractable and give the explicit conditions as matrix inequalities, using the ellipsoidal sets.

Now, we recall the main result of the work [9]. First, some definitions should be given before the main results.

Given the system (1) and $x \in \mathbb{R}^n$ and $\eta \in \mathbb{R}^n$, we define

$$\begin{array}{ll}
\mathcal{N}^+(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} > 0, \; K_i x < -1\},\\
\mathcal{N}^-(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} < 0, \; K_i x > 1\},\\
\mathcal{N}(x,\eta) & = \mathcal{N}^+(x,\eta) \cup \mathcal{N}^-(x,\eta),\\
\mathcal{P}^+(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} > 0, \; K_i x > 1\},\\
\mathcal{P}^-(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} < 0, \; K_i x < -1\},\\
\mathcal{P}(x,\eta) & = \mathcal{P}^+(x,\eta) \cup \mathcal{P}^-(x,\eta),\\
\mathcal{L}^+(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} > 0, \; |K_i x| \le 1\},\\
\mathcal{L}^-(x,\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} < 0, \; |K_i x| \le 1\},\\
\mathcal{L}(x,\eta) & = \mathcal{L}^+(x,\eta) \cup \mathcal{L}^-(x,\eta),\\
\mathcal{O}(\eta) & = \{i \in \mathbb{N}_m : \, \eta^T B_{(i)} = 0\}.
\end{array} \tag{4}$$

The sets defined in (4), subsets of $\mathbb{N}_m$, permit to characterize the regions of $(x, \eta) \in \mathbb{R}^n \times \mathbb{R}^n$ where the input saturates and the different values of $\eta^T B_{(i)}\varphi_i(Kx)$ in these regions. In fact :

$$\left\{\begin{array}{ll}
\eta^T B_{(i)}\varphi_i(Kx) = -\eta^T B_{(i)}, & \text{if } i \in \mathcal{N}^+(x,\eta),\\
\eta^T B_{(i)}\varphi_i(Kx) = \eta^T B_{(i)}, & \text{if } i \in \mathcal{N}^-(x,\eta),\\
\eta^T B_{(i)}\varphi_i(Kx) = \eta^T B_{(i)}, & \text{if } i \in \mathcal{P}^+(x,\eta),\\
\eta^T B_{(i)}\varphi_i(Kx) = -\eta^T B_{(i)}, & \text{if } i \in \mathcal{P}^-(x,\eta),\\
\eta^T B_{(i)}\varphi_i(Kx) = \eta^T B_{(i)} K_i x, & \text{if } i \in \mathcal{L}(x,\eta),\\
\eta^T B_{(i)}\varphi_i(Kx) = 0, & \text{if } i \in \mathcal{O}(\eta),
\end{array}\right. \tag{5}$$

for every $i \in \mathbb{N}_m$. Define an equivalence relation $\sim_x$ on $\mathbb{R}$ denoting $y \sim_x z$ if and only if either: $y > 1$ and $z > 1$; or $y < -1$ and $z < -1$; or $|y| \le 1$ and $|z| \le 1$.

A second equivalence relation $\sim_\eta$ on $\mathbb{R}$ is defined saying that $d, e \in \mathbb{R}$ are such that $d \sim_\eta e$ if and only if either: $d > 0$ and $e > 0$; or $d < 0$ and $e < 0$; or $d = 0$ and $e = 0$. We define two partitions of $\mathbb{R}^n$ induced by the equivalence relations $\sim_x$ and $\sim_\eta$ as

$$\begin{array}{l}
\mathcal{I} = \{J \subseteq \mathbb{R}^n : \; x, \bar{x} \in J \;\Leftrightarrow\; x_i \sim_x \bar{x}_i, \; \forall i \in \mathbb{N}_n\},\\
\mathcal{K} = \{E \subseteq \mathbb{R}^n : \; \eta, \bar{\eta} \in E \;\Leftrightarrow\; \eta_i \sim_\eta \bar{\eta}_i, \; \forall i \in \mathbb{N}_m\}.
\end{array} \tag{6}$$

Given $J \in \mathcal{I}$ and $E \in \mathcal{K}$, for every $x, \bar{x} \in J$ and $\eta, \bar{\eta} \in E$, the sets in (4) are the same. That is $\mathcal{N}^+(x,\eta) = \mathcal{N}^+(x,\bar{\eta}) = \mathcal{N}^+(\bar{x},\bar{\eta}) = \mathcal{N}^+(\bar{x},\eta)$ (analogously, for any other set in (4)). Also relations (5) are the same within $J \times E$, for every $J \in \mathcal{I}$ and $E \in \mathcal{K}$. For every $J \in \mathcal{I}$ and $E \in \mathcal{K}$ and given $x \in J$ and $\eta \in E$, denote, with a slight abuse of notation,

$$\begin{array}{ll}
\mathcal{N}^+(J,E) = \mathcal{N}^+(x,\eta), & \mathcal{N}^-(J,E) = \mathcal{N}^-(x,\eta),\\
\mathcal{P}^+(J,E) = \mathcal{P}^+(x,\eta), & \mathcal{P}^-(J,E) = \mathcal{P}^-(x,\eta),\\
\mathcal{L}^+(J,E) = \mathcal{L}^+(x,\eta), & \mathcal{L}^-(J,E) = \mathcal{L}^-(x,\eta),\\
\mathcal{O}(E) = \mathcal{O}(\eta). &
\end{array} \tag{7}$$

Thus, we introduce the necessary and sufficient condition given in [9], which is the basis of our work in this paper.

Theorem 1 ([9]): Given the system (1), the C-set Ω is such that $\alpha\Omega$ is λ-contractive for every $\alpha \in [0, 1]$ if and only if for every $J \in \mathcal{I}$ and $E \in \mathcal{K}$, there exist $\gamma_i^{J,E} \in \mathbb{R}$ and $\sigma_i^{J,E}(x) \in \mathbb{R}$, with $i \in \mathbb{N}_m$, such that

$$\eta^T Ax + \sum_{i\in\mathbb{N}_m}\gamma_i^{J,E}\eta^T B_{(i)}K_i x + \sigma_i^{J,E}(x)\eta^T B_{(i)} \leq \lambda\phi_{\Omega(x)}(\eta), \tag{8}$$

and

$$\begin{cases} \gamma_i^{J,E} \leq 0, & \sigma_i^{J,E}(x) \geq \gamma_i^{J,E} - 1, & \text{if} i \in \mathcal{N}^+(J,E), \\ \gamma_i^{J,E} \leq 0, & \sigma_i^{J,E}(x) \leq 1 - \gamma_i^{J,E}, & \text{if} i \in \mathcal{N}^-(J,E), \\ \gamma_i^{J,E} \geq 0, & \sigma_i^{J,E}(x) \geq 1 - \gamma_i^{J,E}, & \text{if} i \in \mathcal{P}^+(J,E), \\ \gamma_i^{J,E} \geq 0, & \sigma_i^{J,E}(x) \leq \gamma_i^{J,E} - 1, & \text{if} i \in \mathcal{P}^-(J,E), \\ \gamma_i^{J,E} \in \mathbb{R}, & \begin{cases}\sigma_i^{J,E}(x) \geq \gamma_i^{J,E} - 1, \\ \sigma_i^{J,E}(x) \geq 1 - \gamma_i^{J,E},\end{cases} & \text{if } i \in \mathcal{L}^+(J,E), \\ \gamma_i^{J,E} \in \mathbb{R}, & \begin{cases}\sigma_i^{J,E}(x) \leq \gamma_i^{J,E} - 1, \\ \sigma_i^{J,E}(x) \leq 1 - \gamma_i^{J,E},\end{cases} & \text{if} i \in \mathcal{L}^-(J,E), \\ \gamma_i^{J,E} \in \mathbb{R}, & \sigma_i^{J,E}(x) \in \mathbb{R}, & \text{if } i \in \mathcal{O}(E), \end{cases} \tag{9}$$

hold for all $x \in \Omega$ and all $\eta \in \mathbb{R}^n$.

Proof: The proof of this theorem is given in [9]. ■

The theoretical results given by Theorem 1 might be very complex and hardly manageable for the analysis of stability of saturated systems from the computational point of view. Indeed, since they involve generic nonlinear function $\sigma_i^{J,E}(x)$, nonconvex constraints may result.

To deal with this issue, we propose a relaxation of this problem based on the use of an affine function $\sigma_i^{J,E}(x)$, which brings us back to find sufficient conditions when using the ellipsoidal sets.

IV. MAIN RESULTS

In this section, novel sufficient conditions for the contractivity and invariance of a given ellipsoid are presented. The notion of contractivity is given in Section II. The effectiveness of the conditions given in the Theorem 1 requires that $\sigma_i(x)^{J,E}$ might be nonlinear. For this, we work with $\sigma_i(x)^{J,E}$ as affine relaxation.

Given a positive definite matrix P, define the following ellipsoids

$$\begin{array}{l} \mathcal{E}(P) = \{x \in \mathbb{R}^n : x^T Px \leqslant 1\} \\ \mathcal{E}_x(P) = \{y \in \mathbb{R}^n : y^T Py \leq x^T Px\}, \end{array} \tag{10}$$

and notice that $\mathcal{E}_x(P) = \Psi_{\mathcal{E}(P)}(x)\mathcal{E}(P)$, see [7].

Before giving the main contribution of the paper, one helpful proposition is presented.

Proposition 2: Given the system (1), consider the ellipsoid $\mathcal{E}(P)$, with $P \in \mathbb{R}^{n\times n}$ and $P > 0$. We have that: $\phi_{\mathcal{E}_x(P)}(\eta) = \sqrt{x^T Px}\sqrt{\eta^T P^{-1}\eta}, \quad \forall x \in \mathbb{R}^n, \ \forall \eta \in \mathbb{R}^n.$

Proof: Recall that $\phi_{\mathcal{E}(P)}(\eta) = \sqrt{\eta^T P^{-1}\eta}$, for every $P > 0$ and any $\eta \in \mathbb{R}^n$, see [7]. Then, defining $\tilde{P}(x) = (x^T Px)^{-1}P$, we have

$$\begin{aligned} \phi_{\mathcal{E}_x(P)}(\eta) &= \sup_{y\in\mathcal{E}_x(P)} \eta^T y = \sup_{y^T Py \leq x^T Px} \eta^T y \\ &= \sup_{y^T \tilde{P}(x) y \leq 1} \eta^T y \\ &= \sqrt{\eta^T \tilde{P}(x)^{-1}\eta} = \sqrt{x^T Px}\sqrt{\eta^T P^{-1}\eta}, \end{aligned}$$

which proves the proposition. ■

The main result of the paper is the following theorem. The improvements entailed will be highlighted with an illustrative example comparing it with previous results from the literature.

Theorem 2: Given the system (1), the ellipsoid $\mathcal{E}(P)$ is such that $\alpha\mathcal{E}(P)$ is λ-contractive for every $\alpha \in [0, 1]$, if for every $J \in \mathcal{J}$ and $E \in \mathcal{K}$, there exist $\gamma^{J,E} \in \mathbb{R}^m$, $\alpha^{J,E} \in \mathbb{R}^m$ and $\beta^{J,E} \in \mathbb{R}^{m\times n}$ such that:

$$\begin{cases} \begin{pmatrix} \lambda^2 P & M^{J,E^T} \\ M^{J,E} & P^{-1} \end{pmatrix} \geq 0, \\ B\alpha^{J,E} = 0, \end{cases} \tag{11}$$

where,

$$M^{J,E} = A + \sum_{i\in\mathbb{N}_m} (\gamma_i^{J,E} B_{(i)} K_i + B_{(i)}\beta_i^{J,E}), \tag{12}$$

and,

$$\begin{cases}
\bullet \ \textit{if } i \in \mathcal{N}^+(J,E): \\
\begin{cases} \gamma_i^{J,E} \leq 0, \\ \begin{pmatrix} 1 - \gamma_i^{J,E} + \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (1 - \gamma_i^{J,E} + \alpha_i^{J,E})P \end{pmatrix} \geq 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{N}^-(J,E): \\
\begin{cases} \gamma_i^{J,E} \leq 0, \\ \begin{pmatrix} 1 - \gamma_i^{J,E} - \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (1 - \gamma_i^{J,E} - \alpha_i^{J,E})P \end{pmatrix} \geq 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{P}^+(J,E): \\
\begin{cases} \gamma_i^{J,E} \geq 0, \\ \begin{pmatrix} -1 + \gamma_i^{J,E} + \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (-1 + \gamma_i^{J,E} + \alpha_i^{J,E})P \end{pmatrix} \geq 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{P}^-(J,E): \\
\begin{cases} \gamma_i^{J,E} \geq 0, \\ \begin{pmatrix} -1 + \gamma_i^{J,E} - \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (-1 + \gamma_i^{J,E} - \alpha_i^{J,E})P \end{pmatrix} \geq 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{L}^+(J,E): \\
\begin{cases} \gamma_i^{J,E} \in \mathbb{R}, \\ \begin{pmatrix} 1 - \gamma_i^{J,E} + \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (1 - \gamma_i^{J,E} + \alpha_i^{J,E})P \end{pmatrix} \geqslant 0, \\ \begin{pmatrix} \gamma_i^{J,E} - 1 + \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (\gamma_i^{J,E} - 1 + \alpha_i^{J,E})P \end{pmatrix} \geqslant 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{L}^-(J,E): \\
\begin{cases} \gamma_i^{J,E} \in \mathbb{R}, \\ \begin{pmatrix} 1 - \gamma_i^{J,E} - \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (1 - \gamma_i^{J,E} - \alpha_i^{J,E})P \end{pmatrix} \geqslant 0, \\ \begin{pmatrix} \gamma_i^{J,E} - 1 - \alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (\gamma_i^{J,E} - 1 - \alpha_i^{J,E})P \end{pmatrix} \geqslant 0, \end{cases} \\
\bullet \ \textit{if } i \in \mathcal{O}(E): \\
\gamma_i^{J,E}, \alpha_i^{J,E} \in \mathbb{R}, \ \beta_i^{J,E} \in \mathbb{R}^{1\times n}.
\end{cases} \tag{13}$$

Proof: To prove this theorem, we have to find the formulation of the conditions given by Theorem 1 when using the ellipsoidal sets. To render the problem computationally tractable, the choice of the function $\sigma_i^{J,E}(x)$ is restricted to an affine function. This choice implies that the obtained conditions are sufficient.

First, we substitute $\sigma_i^{J,E}(x)$ with $\beta_i^{J,E}x + \alpha_i^{J,E}$, the equation (8) for $x \in \mathcal{E}(P)$ is equivalent to

$$\eta^T N^{J,E}(x) \leq \lambda \phi_{\mathcal{E}_x(P)}(\eta), \quad \forall x \in \mathcal{E}(P), \ \forall \eta \in \mathbb{R}^n, \tag{14}$$

for every $J \in \mathcal{I}$ and $E \in \mathcal{K}$, with $N^{J,E}(x) = M^{J,E}x + B\alpha^{J,E}$, where $M^{J,E}$ is given by (12).

Now, from Proposition 2, we have that the condition (14) is equivalent to

$$\eta^T N^{J,E}(x) \leq \lambda \sqrt{x^T P x}\sqrt{\eta^T P^{-1}\eta}, \quad \forall x \in \mathcal{E}(P), \ \forall \eta \in \mathbb{R}^n, \tag{15}$$

for every $J \in \mathcal{I}$ and $E \in \mathcal{K}$. In this case, the explicit dependence on η can be removed to obtain a formulation of the condition involving only the state x.

Given $\eta \in \mathbb{R}^n$ with $\eta \neq 0$, define $\hat{\eta} = (\eta^T P^{-1}\eta)^{-1/2}\eta$, and notice that $\hat{\eta} \in \partial\mathcal{E}(P^{-1})$, in fact $\hat{\eta}^T P^{-1}\hat{\eta} = 1$. Thus, apart from the trivial case of $\eta = 0$, (15) is equivalent to

$$\hat{\eta}^T N^{J,E}(x) \leq \lambda\sqrt{x^T P x}, \quad \forall \hat{\eta} \in \partial\mathcal{E}(P^{-1}), \ \forall x \in \mathcal{E}(P),$$

and then, since the supremum of a linear function over a bounded convex set is attained at its boundary, we have

$$\begin{aligned}\sup_{\hat{\eta} \in \partial\mathcal{E}(P^{-1})} N^{J,E^T}(x)\hat{\eta} &= \sup_{\hat{\eta} \in \mathcal{E}(P^{-1})} N^{J,E^T}(x)\hat{\eta} \\ &= \phi_{\mathcal{E}(P^{-1})}(N^{J,E}(x)) \leq \lambda\sqrt{x^T P x}, \quad \forall x \in \mathcal{E}(P).\end{aligned}$$

Thus, from the expression of the support function of $\mathcal{E}(P)$ at $N^{J,E}(x)$ given above, the condition in Theorem 1 is implied by the following condition:

$$N^{J,E^T}(x) P N^{J,E}(x) \leq \lambda^2 x^T P x, \quad \forall x \in \mathcal{E}(P), \tag{16}$$

for every $J \in \mathcal{I}$ and $E \in \mathcal{K}$.

We define for each $J \in \mathcal{I}$ and $E \in \mathcal{K}$ the set
$\mathcal{E}^{J,E}(P) = \{x \in \mathbb{R}^n : N^{J,E^T}(x) P N^{J,E}(x) \leq \lambda^2 x^T P x\}$.

Notice that (16) holds if and only if $x \in \mathcal{E}(P)$, i.e. $x^T P x \leq 1$, implies $x \in \mathcal{E}^{J,E}(P)$. Then the condition (16) is equivalent to:

$$\mathcal{E}(P) \subseteq \mathcal{E}^{J,E}(P). \tag{17}$$

According to this equivalence, $\alpha\mathcal{E}(P)$ is λ-contractive for all $\alpha \in [0, 1]$ if the ellipsoid $\mathcal{E}(P)$ lies in the intersection of the ellipsoids $\mathcal{E}^{J,E}(P)$, for every $J \in \mathcal{I}$, and $E \in \mathcal{K}$.

To transform the condition (17) to matrix inequality, we rewrite the ellipsoid $\mathcal{E}^{J,E}(P)$ as in the following
$\mathcal{E}^{J,E}(P) = \{x \in \mathbb{R}^n : x^T A^{J,E} x + 2b^{J,E^T} x + c^{J,E} \leq 0\}$, with, $A^{J,E} = M^{J,E^T} P M^{J,E} - \lambda^2 P$,
$b^{J,E^T} = \alpha^{J,E^T} B^T P M^{J,E}$ and $c^{J,E} = \alpha^{J,E^T} B^T P B \alpha^{J,E}$. We also, describe $\mathcal{E}(P)$ as $\mathcal{E}(P) = \{Du + d : \|u\|_2 \leq 1\}$, with $D = P^{-1/2}$ and $d = 0$.

Thus, using the last notation of the ellipsoids, the condition (17) holds if and only if:

$$\begin{aligned}&\sup_{\|u\|_2 \leq 1} (u^T D^T A^{J,E} D u + 2b^{J,E^T} D u + c^{J,E}) \\ &= c^{J,E} + \sup_{\|u\|_2 \leq 1} (u^T D^T A^{J,E} D u + 2b^{J,E^T} D u) \leq 0,\end{aligned}$$

for more details see [7]. This last inequality is satisfied if and only if :

$$\sup_{\|u\|_2 \leq 1} (u^T D^T A^{J,E} D u + 2b^{J,E^T} D u) \leq -c^{J,E}. \tag{18}$$

Using the S-procedure, we prove that the condition (18) is true if and only if there exists a $\delta^{J,E} \geq 0$ such that:

$$\begin{pmatrix} -\delta^{J,E} - c^{J,E} & b^{J,E^T} D^T \\ D b^{J,E} & \delta^{J,E} I_n - D A^{J,E} D \end{pmatrix} \geq 0.$$

Replacing the parameters $A^{J,E}$, $b^{J,E}$, $c^{J,E}$ and D with their expressions, we find

$$\begin{pmatrix} -\delta^{J,E} - \alpha^{J,E^T} B^T P B \alpha^{J,E} & (P^{-1/2} M^{J,E^T} P B \alpha^{J,E})^T \\ P^{-1/2} M^{J,E^T} P B \alpha^{J,E} & (\delta^{J,E} + \lambda^2) I_n - P^{-1/2} M^{J,E^T} P M^{J,E} P^{-1/2} \end{pmatrix} \geq 0. \tag{19}$$

For the inequality (19) to be satisfied, it is necessary that $-\delta^{J,E} - \alpha^{J,E^T} B^T P B \alpha^{J,E} \geq 0$, which implies that $\delta^{J,E} = 0$ and $B\alpha^{J,E} = 0$. Then, the inequality (19) is equivalent to

$$\begin{cases} (\delta^{J,E} + \lambda^2) I_n - P^{-1/2} M^{J,E^T} P M^{J,E} P^{-1/2} \geq 0, \\ B\alpha^{J,E} = 0, \\ \delta^{J,E} = 0. \end{cases}$$

which is equivalent also to

$$\begin{cases} \lambda^2 P - M^{J,E^T} P M^{J,E} \geq 0, \\ B\alpha^{J,E} = 0, \\ \delta^{J,E} = 0. \end{cases}$$

Now, using the Schur complement, see [3], the latter inequalities are equivalent to

$$\begin{cases} \begin{pmatrix} \lambda^2 P & M^{J,E^T} \\ M^{J,E} & P^{-1} \end{pmatrix} \geq 0, \\ B\alpha^{J,E} = 0, \\ \delta^{J,E} = 0, \end{cases}$$

which is equivalent to (11).

Thus, the condition (11) and (14) are equivalent, for each $J \in \mathcal{I}$ and $E \in \mathcal{K}$. Therefore, the condition (11) and (8) are also equivalent, for each $J \in \mathcal{I}$ and $E \in \mathcal{K}$.

For proving that (13) implies (9), we will use the definition of the support function and we have to find the equivalent of the conditions (9) in the ellipsoidal case. The conditions in (9) are similar in the form. Thus, just the first condition will be dealt with.

Taking the case when $i \in \mathcal{N}^+(J, E)$ and substituting $\sigma_i^{J,E}(x)$ with $\beta_i^{J,E}x + \alpha_i^{J,E}$, the condition given in Theorem 1 becomes

$$\gamma_i^{J,E} \leq 0, \qquad \beta_i^{J,E} x + \alpha_i^{J,E} \geq \gamma_i^{J,E} - 1, \qquad \forall\, x \in \mathcal{E}(P),$$

that is,

$$\gamma_i^{J,E} \leq 0, \qquad -\beta_i^{J,E} x \leq 1 - \gamma_i^{J,E} + \alpha_i^{J,E}, \qquad \forall\, x \in \mathcal{E}(P).$$

Using the definition of the support function, we find that it is equivalent also to:

$$\gamma_i^{J,E} \leq 0, \qquad \beta_i^{J,E} P^{-1} \beta_i^{J,E^T} \leq (1 - \gamma_i^{J,E} + \alpha_i^{J,E})^2.$$

Now, applying the Schur complement, we end up with

$$\left\{ \begin{array}{l} \gamma_i^{J,E} \leq 0, \\ \begin{pmatrix} 1-\gamma_i^{J,E}+\alpha_i^{J,E} & \beta_i^{J,E} \\ \beta_i^{J,E^T} & (1-\gamma_i^{J,E}+\alpha_i^{J,E})P \end{pmatrix} \geq 0. \end{array} \right. \tag{20}$$

Thus, (20) is equivalent to the condition (9) when $i \in \mathcal{N}^+(J, E)$, $\Omega = \mathcal{E}(P)$ and $\sigma^{J,E}$ is an affine function. For the conditions in the other regions, the same reasoning leads to the results given in Theorem 2.

Thus, the conditions given in Theorem 2 are proved. ■

Remark 1: Notice that, since we imposed a structure for $\sigma^{J,E}$, that is affinity with respect to x, the condition given in Theorem 2 is just sufficient for λ-contractivity of $\alpha\mathcal{E}(P)$ for all $\alpha \in [0, 1]$. Nevertheless, our opinion is that it might be also necessary for the ellipsoidal case, as our numerical tests seem to infer. This issue will be studied in a future work.

Although the theory is developed in [9] also for asymmetric saturations, symmetry should be maintained, in our opinion, when dealing with ellipsoids to avoid overly conservative results. Considering polytopic sets, thus potentially asymmetric, for asymmetric saturations is one of our next research directions

In order to illustrate the effectiveness of the results presented in this work, some computational considerations and an illustrative example will be given in the next section.

V. COMPUTATION AND ILLUSTRATIVE EXAMPLE

A. Computational considerations

Theorem 2 can be applied to the computation of the maximal ellipsoidal set, as estimate of the domain of attraction of a saturated system. Nevertheless, since the sufficient conditions given in Theorem 2 are nonlinear matrix inequalities, then the direct application of this theorem might be computationally demanding.

It can be noted about the conditions given in Theorem 2, that all the nonlinear terms depend on P and then for a given P the conditions become LMIs. Therefore, the Theorem 2 permits to determine the maximal ρ, such that $\rho\mathcal{E}(P)$ is contractive by using dichotomy. This technique is proposed in the following algorithm.

Algorithm 1 : Computing the contractively invariant ellipsoid for discrete-time saturated system (1).

1: Given the saturated discrete-time linear system (1).
2: Find the shape of an initial λ-contractive ellipsoidal set $\mathcal{E}(P)$, using the existing approaches, and consider the resulting matrix P.
3: Replace P by $(\rho)^{-1}P$ in Theorem 2.
4: **Solve** : $\rho_{max} = \max\limits_{\rho \geq 1} \rho$ **s.t.** the conditions (11) and (13) hold, using the dichotomy optimization.
5: The result is $\rho_{max}\mathcal{E}(P) = \{x \in \mathbb{R}^n : x^T P x \leqslant \rho_{max}\}$.

Algorithm 1 permits to compute the ρ_{max}, which is the factor used to inflate the initial λ-contractive ellipsoid $\mathcal{E}(P)$ to get the maximal λ-contractive ellipsoid using the dichotomy method. The results given by this algorithm are obtained for a given matrix P. However, considering P fixed is not very restrictive, see the work carried out in [14].

B. Illustrative example

In order to highlight the improvements provided by the proposed method, a comparison with some equivalent results from the literature is given, see [1]. One illustrative example will be presented.

First, we recall the result presented in [1], for the continuous-time context but extendable to the discrete-time one, which was used in [9].

Theorem 3: Given the system (1), and the ellipsoid $\mathcal{E}(P)$, with $P \in \mathbb{R}^{n\times n}$ and $P > 0$, if for every $J \subseteq \mathbb{N}_m$ and every $i \in J$, there exists $G_i^J \in \mathbb{R}^{1\times n}$ such that

$$\begin{array}{ll} (N^J)^T P N^J \leq \lambda P, & \forall J \subseteq \mathbb{N}_m, \\ G_i^J P^{-1} (G_i^J)^T \leq 1, & \forall J \subseteq \mathbb{N}_m,\ \forall i \in J, \end{array}$$

where $N^J = A + \sum_{i\in\bar{J}} B_{(i)} K_i + \sum_{i\in J} B_{(i)} G_i^J$, then $\alpha\mathcal{E}$ is λ-contractive, with $\lambda \in [0, 1]$, for every $\alpha \in [0,\ 1]$.
The quadratic stability conditions for saturated systems given in [1], [13], [15] are substantially based on Theorem 3.

The example considered below is the one presented in [9], increasing its size. The authors in [9] demonstrate analytically, that the system is λ-contractive for all $x \in \mathbb{R}$. We choose this example to show that we can obtain the exact, analytical solution by applying the convex-optimization based Algorithm 1. We also show that the obtained result is less conservative than those found in the literature.

Example 1 ([9]): Consider the two-dimensional system with two saturation inputs:

$$\left\{ \begin{array}{l} x_1^+ = x_1 + 0.5\varphi(x_1) - \varphi(0.5x_1), \\ x_2^+ = 0.5x_2. \end{array} \right. \tag{21}$$

that is (1), with $n = 2$, $m = 2$ and $A = \begin{pmatrix} 1 & 0 \\ 0 & 0.5 \end{pmatrix}$, $B = \begin{pmatrix} 0.5 & -1 \\ 0 & 0 \end{pmatrix}$, $K = \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix}$, and the ellipsoid $\mathcal{E}(P) \subseteq R^2$. Consider $\lambda = 1$. We have that

$$f(x) = \begin{pmatrix} x_1^+ \\ x_2^+ \end{pmatrix} = \left\{ \begin{array}{ll} \begin{pmatrix} x_1 \\ 0.5x_2 \end{pmatrix}, & \text{if } x_1 \in [0,\ 1), \\ \begin{pmatrix} 0.5x_1 + 0.5 \\ 0.5x_2 \end{pmatrix}, & \text{if } x_1 \in [1,\ 2), \\ \begin{pmatrix} 0.5x_1 - 0.5 \\ 0.5x_2 \end{pmatrix}, & \text{if } x_1 \in [2,\ \infty). \end{array} \right.$$

We remark that in this system, there is not a region of x at which the second input is saturated but not the first one. However, the other regions exist and in Theorem 3 all the possible combinations on saturations are implicitly considered. Our method overcomes this situation and get a better result even when we take the combinations which do not exist. We also note that the system in the Example 1 is globally stable in $\mathbb{R}^2$, see [9] and that invariance of $\mathcal{E}(I) = \{x \in \mathbb{R}^n : x^T I x \leq \rho\}$ is ensured for all $\rho \geq 0$.

We use different methods to estimate the domain of attraction for this system by choosing the reference set as $\Gamma_R = co\{x_1, x_2\}$, where $x_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$ and $x_2 = \begin{pmatrix} -1 & 0 \end{pmatrix}^T$. The objective is to compute the maximal value of τ such that $\tau\Gamma_R \subseteq \mathcal{E}(P)$ for some $P \geq 0$ by using Theorem 3 and Theorem 2. The initial λ-contractive ellipsoidal set $\mathcal{E}(P)$ used in the Algorithm 1 is given by Theorem 3.

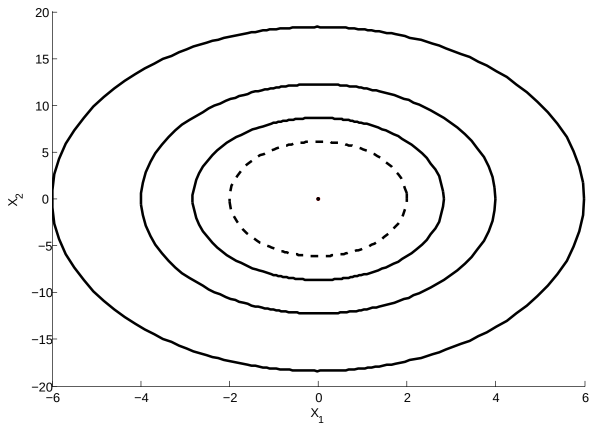

Fig. 1. Contractive ellipsoidal set obtained with Theorem 3 (dashed line) and Theorem 2 (solid line)

Applying Theorem 3 to the system (21) subject to the objective given above, we get as solution $P^* = \begin{pmatrix} 0.2500 & 0.0000 \\ 0.0000 & 0.0088 \end{pmatrix}$, that is used in Algorithm 1. The ellipsoidal estimation set of the domain of attraction given by Theorem 3 is showed in Figure 1 with a dashed line. Now, using Algorithm 1, the result is that $\rho\mathcal{E}(P^*)$ is invariant for all $\rho \geq 0$. Figure 1 shows some contractively invariant ellipsoidal sets obtained using Algorithm 1 with a solid line. This is the exact result obtained analytically in [9].

Figure 1 shows that the ellipsoids obtained by the sufficient condition presented in this paper are greater than that determined by Theorem 3. We remark that in the x_1 axis, the limits of the ellipsoid given by Theorem 3 do not exceed 2 and -2, which are the boundary of the region where the second input is saturated but not the first one, inadmissible as notice before. However, when using the result presented in this paper, the set $\rho\mathcal{E}(P^*)$ is proved to be invariant for all $\rho \in \mathbb{R}$.

The fact is that in Theorem 3, all possible combinations on saturations are implicitly considered even those that do not exist and this affects the result. Thus, according to the result given in Figure 1, this is an important source of conservatism, overcome by our method.

Notice that, the numerical application of the sufficient condition given in this paper suffer from a complexity that is exponential in the dimension of the control inputs. However, the computational complexity is not affected by the system dimension, that can be high-dimensional, provided the number of input is small.

VI. CONCLUSION

This paper examined the problem of stability and convergence for a discrete-time linear systems subject to actuator saturation by determining the contractively invariant ellipsoidal sets. Sufficient condition for an ellipsoid to be invariant and contractive for saturated system has been given. The method determines the local quadratic Lyapunov function. It was demonstrated that the proposed approach is less conservative then those found in the literature. Some computational considerations was provided. An algorithm for the characterization of the contractively invariant ellipsoid was given. An illustrative example was presented to elucidate the main contribution of the paper.

REFERENCES

[1] T. Alamo, A. Cepeda, and D. Limon. Improved computation of ellipsoidal invariant sets for saturated control systems. In *Proc. of the 44th IEEE Conf. on Decision and Control and European Control Conf. 2005, CDC-ECC'05*, pages 6216–6221, Seville, Spain, Dec. 2005.

[2] T. Alamo, A. Cepeda, D. Limon, and E. F. Camacho. A new concept of invariance for saturated systems. *Automatica*, 42:1515–1521, 2006.

[3] T. Alamo, J. E. Normey-Rico, M. R. Arahal, D. Limon, and E. F. Camacho. Introducing linear matrix inequalities in a control course. *IFAC*, 2006.

[4] G. Duan B. Zhou and Z. Lin. Approximation and monotonicity of the maximal invariant ellipsoid for discrete-time systems by bounded controls. *IEEE Transactions on Automatic Control*, 55(2), Feb. 2010.

[5] F. Blanchini. Set invariance in control. *Automatica*, 35:1747–1767, 1999.

[6] F. Blanchini and S. Miani. *Set-Theoretic Methods in Control*. Birkhäuser, 2008.

[7] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in system and control theory*. SIAM, 1994.

[8] M. Fiacchini, S. Tarbouriech, and C. Prieur. Invariance of symmetric convex sets for discrete-time saturated systems. In *Proc. of the 50th IEEE Conf. on Decision and Control and European Control Conf. 2011, CDC-ECC'11*, pages 7343–7348, Orlando, FL, Dec. 2011.

[9] M. Fiacchini, S. Tarbouriech, and C. Prieur. Necessary and sufficient conditions for invariance and convex sets for discrete-time saturated systems. In *Proc. of the 52th IEEE Conf. on Decision and Control (CDC), 2013*, pages 3788–3793, Florence, Italy, Dec. 2013.

[10] E. G. Gilbert and K. Tan. Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, 36:1008–1020, 1991.

[11] J. M. Gomes da Silva Jr. and S. Tarbouriech. Local stabilization of discrete-time linear systems with saturating controls: An lmi-based approach. *IEEE Transactions on Automatic Control*, 46(1), 2001.

[12] J. M. Gomes da Silva Jr. and S. Tarbouriech. Antiwindup design with guaranteed regions of stability: An lmi-based approach. *IEEE Transactions on Automatic Control*, 50(1), Jan. 2005.

[13] T. Hu and Z. Lin. Exact characterization of invariant ellipsoids for single input linear systems subject to actuator saturation. *IEEE Transactions on Automatic Control*, 47(1):164–169, 2002.

[14] T. Hu and Z. Lin. Composite quadratic lyapunov functions for constrained control systems. *IEEE Transactions on Automatic Control*, 48(3):440–452, Dec. 2003.

[15] T. Hu, T. Thibodeau, and A. R. Teel. A unified Lyapunov approach to analysis of oscillations and stability for systems with piecewise linear elements. *IEEE Transactions on Automatic Control*, 55(12):2864–2869, 2010.

[16] B. E. A. Milani. Piecewise-affine lyapunov functions for discrete-time linear systems with saturating controls. *Automatica*, 38:2177–2184, 2002.

[17] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, USA, 1970.

[18] S. Tarbouriech, G. Garcia J. M. Gomes da Silva Jr., and I. Queinnec. *Stability and Stabilization of Linear Systems with Saturating Actuators*. Springer, 2011.

[19] L. Zaccarian and A. R. Teel. *Modern Anti-windup Synthesis: Control Augmentation for Actuator Saturation*. Princeton Series in Applied Mathematics. Princeton University Press, 2011.

[20] B. Zhou, W. X. Zheng, and G.-R. Duan. Stability and stabilization of discrete-time periodic linear systems with actuator saturation. *Automatica*, 47(47):1813–1820, 2011.